

**How We Can Change Your Mind:
Anodal tDCS to Fp3 alters human stimulus representation and learning.**

Ciro Civile, Rossy McLaren, and I.P.L. McLaren

School of Psychology, College of Life and Environmental Sciences,

University of Exeter, UK

Correspondence should be addressed to either **Ciro Civile** or **Ian McLaren**

School of Psychology, College of Life and Environmental Sciences,

University of Exeter, UK.

E-mails: c.civile@exeter.ac.uk, i.p.l.mclaren@exeter.ac.uk

Please cite as: Civile, C., McLaren, R.P., and McLaren, I.P.L., (2018). How We Can Change Your Mind: Anodal tDCS to Fp3 alters human stimulus representation and learning.

Neuropsychologia, doi.org/10.1016/j.neuropsychologia.2018.08.019

Abstract

The aim of the current work is to advance our understanding of both the mechanisms controlling perceptual learning and the face inversion effect. In the three double blind experiments reported here (total N=144) we have shown that anodal tDCS stimulation (10 mins at 1.5mA) delivered over the left DLPFC at Fp3 affects perceptual learning and drastically reduces the, usually robust, face inversion effect. In Experiment 1, we found a significantly reduced inversion effect in the anodal group compared to that in the sham group. Experiment 2 replicated the pattern of results found in Experiment 1. In both experiments recognition performance for upright faces in the anodal group was significantly impaired compared to that in the sham group. Finally, using an *active control* in Experiment 3 (same behavioural task but different tDCS targeted brain area) we showed that the same Fp3 anodal tDCS stimulation effect is not obtained when a different brain area is targeted.

Key Words: Perceptual Learning, tDCS, Face Inversion Effect; Generalization

Perceptual learning generally denotes an improvement in learning with experience, such that familiarisation with the stimuli to be discriminated aids later performance (Gold & Watanabe, 2010). In this paper we will look at a particular version of perceptual learning, in which familiarisation with a class of stimuli promotes later discrimination among members of that class, even though the stimuli used at test are not those used for familiarisation.

Although, perceptual learning of this type is well established as a phenomenon (see McLaren, Leever and Mackintosh (1994) for one of the earliest experimental examples in humans), and we have some idea how it works, we do not know as much as we would like about the neural mechanisms involved or how to influence it. In this paper we set out to show that a particular form of transcranial Direct Current Stimulation (tDCS) can influence perceptual learning, causing a change in the way that people process images of faces (and other stimuli) so as to significantly reduce the inversion effect that would otherwise occur in an old/new recognition paradigm.

The face inversion effect refers to a processing disadvantage for upside-down (i.e. inverted) face images compared to upright ones. In 1969, when it was first discovered by Yin (and subsequently reported by many others) the inversion effect was used as a robust marker for “special” face processing. This because studies reported the FIE to be larger for face stimuli than for other visual stimuli (e.g. house) (Valentine & Bruce, 1988; Yovel & Kanwisher, 2005). Diamond and Carey (1986)’s finding of an inversion effect for dog images when participants were dog breeders (vs novices), and Gauthier’s work on perceptual expertise and the inversion effect for novel objects named Greebles (Gauthier & Tarr, 1997) challenged the idea that faces are special and highlighted “expertise” as a contributing factor to the inversion effect. At the same time McLaren (1997)’s work made the first case for *perceptual learning* playing a vital role in the face inversion effect. Specifically, it was demonstrated that after a short period of *pre-exposure* to checkerboard exemplars drawn from

a prototype-defined category, participants would benefit when learning to discriminate between a new pair of upright exemplars from that now familiar category, but not if a pair of inverted exemplars from that same category were involved. These findings were later extended to an *old/new recognition* paradigm of the type conventionally used in studies that demonstrate the face inversion effect (Civile, Zhao, Ku, Elchhleppe, Lavric & McLaren, 2014). Taken together, these studies provide evidence that exposure to, or experience with a set of stimuli drawn from a prototype defined category can improve within category discrimination, and that this in turn leads to a type of inversion effect with artificial stimuli that is similar to that found in faces.

Our explanation of the inversion effect for checkerboards has as its basis the theory of representation development put forward by McLaren, Kaye, & Mackintosh (1989). More specifically, it relies on the differential latent inhibition of common elements mechanism first outlined in that paper. According to this theory, pre-exposure helps because it results in the unique elements or features of a stimulus (which help us discriminate between stimuli) being favoured (i.e. relatively more active) during learning compared to the common elements or features shared by the stimuli (which do not help in discrimination, they promote generalisation). This is a consequence of the common elements suffering from greater salience reduction than the unique elements as a result of latent inhibition because they are both more predictable and more frequently encountered, and so develop stronger associations both between themselves and from other elements present (McLaren & Mackintosh, 2000; McLaren, Forrest & McLaren 2012) resulting in a reduction in error, and a consequent loss of salience/activation.

It may not seem immediately obvious how the effect just considered, latent inhibition, could lead to perceptual learning, but the key is in the stimulus representations that are involved. If, say, we have two similar stimuli, A and B, that share many features, then the

elements representing these features (common elements) will tend to be the source of any confusion when we try and discriminate one stimulus from the other. Pre-exposure to these stimuli will have the effect of decreasing the salience of these elements relative to others that are unique to one of the stimuli. This is partly because they are encountered twice as often – so associations between them form rapidly – and also because they are highly predictive of one another and reliably predicted by other elements present. The unique elements of A, however, only occur in A (by definition), and so can only be predicted by the common elements some of the time (i.e. when A occurs rather than B). Conversely, when A is presented, the unique A elements can always predict the common elements, and a similar analysis applies to presentations of B. The result is that the common elements, which we will denote by x , would decline in salience rapidly. The unique A elements (a) and unique B elements (b) remain relatively salient, and it is these elements that allow discrimination between A and B, hence perceptual learning. The advantage gained in discriminability outweighs any loss in overall rate of learning, and, in fact, there may not be any loss in overall rate of learning in any case in humans if changes in attention compensate for the average loss of salience (see McLaren, Graham and Wills, 2010 for a detailed exposition).

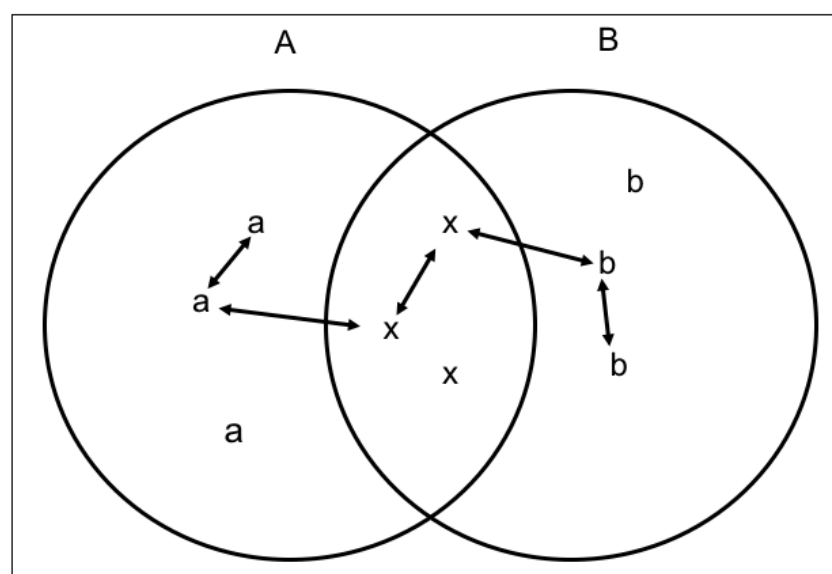


Figure 1. Associations between elements representing features leads to a reduction in salience for those elements, which means less learning to well-predicted features.

Importantly, in terms of the connection with face recognition, Civile et al. (2014)'s findings were also supported by event-related potentials (ERPs) showing a larger and delayed N170 (a negative deflection that occurs 170ms after a face is presented) for inverted checkerboards compared to that for upright checkerboards, both drawn from a familiar prototype-defined category (Civile et al., 2014, Experiment 4). This effect is similar to that usually elicited by upside down faces (Eimer, 2000; Carmel & Bentin, 2002; Zion-Golumbic & Bentin, 2007; Civile, Elchlepp, McLaren, Galang, Lavric & McLaren, 2018),

In a recent study, Civile, Verbruggen, McLaren, Zhao, Ku and McLaren (2016) extended their investigation of the inversion effect for checkerboards (taking this as an index of perceptual learning) by testing the effects of tDCS on it. The authors adopted the same tDCS montage previously used by Ambrus et al (2011) to influence categorisation of prototype-defined stimuli. Ambrus et al (2011) found that anodal tDCS delivered to dorsolateral prefrontal cortex (DLPFC) at the Fp3 site during the training phase of a categorisation task where participants had to identify prototype and low-distortion patterns as category members reduced classification accuracy on test for the prototype compared to sham stimulation. As Civile et al. (2014) obtained the inversion effect using prototype-defined categories of checkerboards, and the MKM model would imply that a strong representation of the prototype is a prerequisite for perceptual learning of the type we are considering, Civile et al. (2016) adopted the same tDCS montage as that used by Ambrus et al. (2011). Civile et al. (2016) were able to provide some evidence that anodal tDCS at the Fp3 area significantly reduced the inversion effect for checkerboards that was otherwise obtained in the sham group, and did this by affecting performance for upright checkerboards from the familiar category. To be specific, the upright checkerboards taken from a familiar category

were less well recognized than those drawn from the novel category, leading to a reduction of the inversion effect. This remarkable and informative result suggested that perceptual learning in humans can be affected by relatively brief tDCS stimulation.

In line with Civile et al (2016)'s study on tDCS and perceptual learning for checkerboards and some more recent pilot work with faces (Civile, Obhi, & McLaren, 2018), in the current paper we report the results from three experiments that establish how the tDCS montage first used by Ambrus et al (2011) to reduce the prototype distortion effect, when extended to perceptual learning can significantly affect the robust inversion effect that is usually found for faces.

Hence, in Experiment 1 we delivered tDCS anodal stimulation at Fp3 site and sham as a control (between-subjects) while participants performed an old/new recognition task with face stimuli. Experiment 2 aimed to replicate and thus establish the effects found in Experiment 1. Finally, Experiment 3 is an *active control* experiment that uses the same behavioural task but targeted a different tDCS brain area to test if we can still obtain the effects found in Experiment 1 and 2. In the discussion we compare the results we obtained in these experiments to those from the Civile et al (2016) study. We will conclude that our results require an explanation in terms of a salience reduction account of perceptual learning based on the MKM model.

Method

Subjects

Overall, 144 naïve subjects (111 women; mean age = 21, age range = 18-59 years) took part in the three experiments. Each experiment included 48 subjects randomly assigned to either sham or anodal tDCS groups (24 in each group). All the subjects were from the University of Exeter (mostly students) and were given course credit or cash for their participation. They were right-handed and were selected according to the safety screening criteria approved by the Research Ethics Committee at the University of Exeter. The sample

size was determined from earlier studies that used the same behavioural paradigm (i.e., an old/new recognition task) for face stimuli (Civile, McLaren & McLaren, 2014; 2016) and tDCS montage (Civile, Obhi and McLaren, 2018).

Materials

The study used a set of 128 male and 128 female face images standardized to grayscale on a black background (pics.stir.ac.uk). These face stimuli were the same as those used in previous studies with the old/new recognition task paradigm (Civile, McLaren, & McLaren, 2014, 2016). The stimuli, whose dimensions were 5.63 cm x 7.84 cm, were presented at a resolution of 1280 x 960 pixels. The experiment was run using Superlab 4.0.7b. on an iMac computer. Participants sat about 70 cm away from the screen on which the images were presented.

Behavioral Task

The old/new recognition task consisted of two parts: a ‘study phase’ and an ‘old/new recognition phase’. In the study phase, each subject was shown 64 upright and 64 inverted male and female face stimuli for 128 images in total. The face stimuli were presented one at a time in random order with no response required from the subject. In the old/new recognition phase, 128 novel face stimuli (half upright and half inverted) were added to the 128 face stimuli seen in the study phase, and all 256 images were presented one at a time in random order. The subject had to respond according to whether or not they thought they had seen the face stimuli during the study phase. For a given subject, each face stimulus only appeared in one orientation (upright or inverted) during the experiment.

Following the instructions, in each trial of the study phase subjects saw a fixation cross in the center of the screen, presented for 1 second, then a face stimulus was presented on screen for 3 seconds before moving on to the next trial. After all the 128 face stimuli had been presented, the program displayed another set of instructions, explaining the recognition task. In this task, subjects were asked to press the ‘.’ key if they recognized the face stimulus

as having been shown in the study phase on any given trial, or press 'x' if they did not (the keys were counterbalanced). During the recognition task, the face stimuli were each shown for 3 seconds during which time subjects had to respond (Figure 1a).

tDCS apparatus

Stimulation was delivered by a battery driven constant current stimulator (neuroConn DC-Stimulator Plus) using a pair of surface sponge electrodes (7cm x 5cm i.e. 35 cm²) soaked in saline solution and applied to the scalp at the target areas of stimulation. We adopted a bilateral bipolar-non-balanced montage with one of the electrodes (anode) placed over the target stimulation area (Fp3 or rIFG) and the other (cathode) on the forehead over the reference area (right or left eyebrow respectively). For the Fp3 montage (Experiment 1&2) once we had identified the Cz for each of the subjects (half the distance between the inion and nasion area) we measured 7 cm anterior relative to the Cz and 9cm to the left (see Ambrus et al., 2011, Civile et al., 2016). The right-Inferior Frontal Gyrus (rIFG) has been implicated in several tasks (e.g. go/no go tasks) and previous studies have shown tDCS administered over the rIFG to be effective (for examples on go/ no go tasks see Cunillera, Brignani, Cucurell, Fuentamilla, Miniussi, 2014, 2016; Jacobson, Jawitt, Lavidor, 2011; Stramaccia, Penolazzi, Sartori, Braga, Mondini, Galfano, 2014). However, so far there have been no experiments that looked at the effects of tDCS delivered over the rIFG while participants perform in a perceptual learning task. Thus, we selected the rIFG as the targeted area for our active control study (Experiment 3). To locate the stimulation area for rIFG we adopted what other studies have used before (Cunillera et al., 2014, 2016; Jacobson et al., 2011; Stramaccia et al., 2014). Hence, the rIFG was identified as the area underlying the crossing point between T4-Fz and F8-Cz in the 10-20 EEG system. We measured the heads of several individuals to find the position for the anode between T4-Fz and F8-Cz (EEG equivalent site FT8). We found that, on average, this site can be located by finding Cz (the

mid-point between theinion and nasion area) and then measuring at right-angles (in line with the ear) 14cm and then forwards 3.5cm. The reference electrode was positioned just above the left eyebrow.

All three experiments were conducted using a double-blind procedure reliant on the neuroConn study mode in which the experimenter inputs numerical codes (provided by another experimenter otherwise unconnected with running the experiment), that switch the stimulation mode between “normal” (i.e. anodal) and “sham” stimulation. In the anodal condition, a direct current stimulation of 1.5mA was delivered for 10 mins (5 s fade-in and 5 s fade-out) starting as soon as the subjects began the behavioural task and continuing throughout the study phase only. In the sham group, the identical stimulation mode was displayed on the stimulator and subjects experienced the same 5 s fade-in and 5 s fade-out, but with the stimulation intensity of 1.5mA delivered for just 30 s, following which a small current pulse (3 ms peak) was delivered every 550 ms (0.1mA over 15 ms) for the remainder of the 10 mins to check impedance levels. For every subject the stimulation finished towards the end of the study phase. Although no stimulation took place during the remaining part of the experiment, the tDCS electrodes remained on the participant (Figure 1b & c).

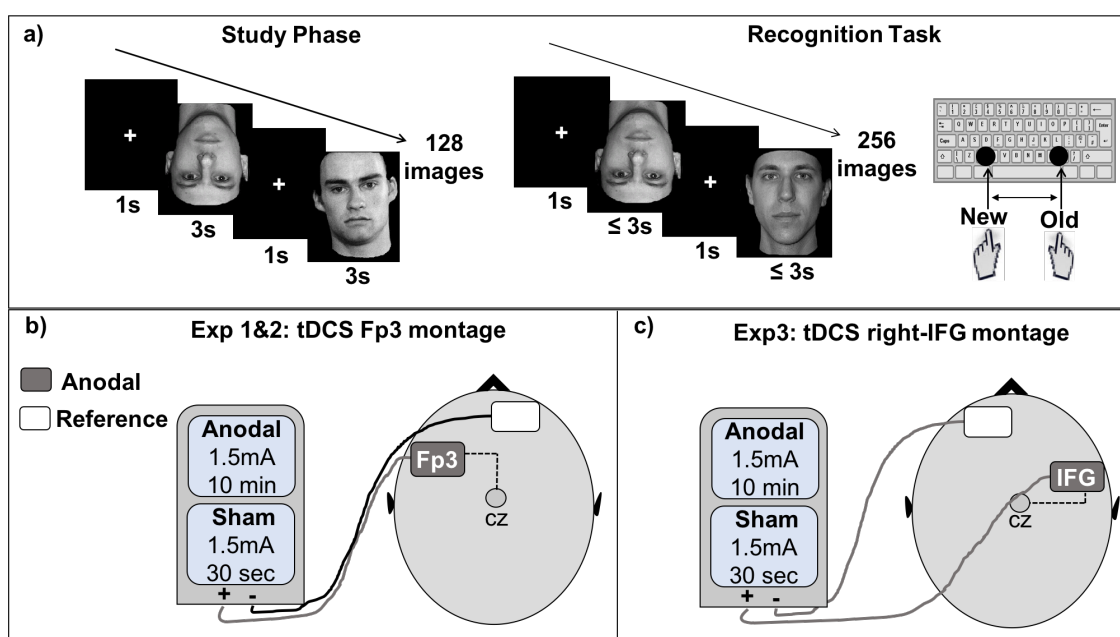


Figure 1. Panel a illustrates the *old/new recognition* task used in the three experiments here reported. **Panel b** shows the tDCS montage adopted in Experiment 1 and 2. This was the same montage used in Civile et al (2016)'s study. **Panel c** shows the tDCS montage used for the *active control* study.

Results

Data Analysis

Our primary measure was performance accuracy in the recognition task. The data from all the participants was used in the signal detection d' analysis of the recognition task (where a $d' = 0$ indicates chance-level performance). We assessed performance against chance to show that both upright and inverted face stimuli in both the tDCS sham and anodal groups across the three experiments were recognized significantly above chance. (For all four conditions we found $p < .005$ for this analysis). Each p-value reported for the comparisons between conditions is two-tailed, and we also report the F or t value along with effect size. We also analysed the data for the accuracy scores which confirmed our results, and for the reaction times (RTs) to check for any speed accuracy trade-off. We do not report these analyses here because they do not add anything to the interpretation of our results.

We computed a 2 x 2 x 3 mixed model design using, as a within-subjects factor, *Face Orientation* (upright or inverted), and the between-subjects factors *tDCS Stimulation* (sham or anodal) and *Experiment* (Fp3, replication, rIFG). A mixed model Analysis of Variance (ANOVA) revealed a significant main effect of *Face Orientation* $F(1, 138) = 146.01, p < .001, \eta^2_p = .51$ as well as a significant interaction between *Face Orientation* and *tDCS Stimulation* $F(1, 138) = 6.16, p = .011, \eta^2_p = .04$. No significant interaction was found between *Face Orientation* and *Experiment* $F(2, 138) = 0.37, p = .688, \eta^2_p = .005$. Importantly, a significant three-way interaction was found, $F(2, 138) = 4.16, p = .018, \eta^2_p = .05$. We decomposed the three-way interaction by examining the two-way interactions (*Face Orientation* x *Stimulation*) separately for the three experiments. Follow-up paired t test

analyses were conducted to compare performance on upright and inverted face stimuli (the inversion effect) in each tDCS group (sham, anodal) within each experiment.

Planned comparisons to measure the inversion effect were used because we have a considerable amount of data (Civile, McLaren, McLaren, 2012, 2014, 2016) showing an inversion effect obtained with these same stimuli and behavioural procedure. Hence, our primary measure is the face inversion effect given by comparing performance on upright and inverted faces. Importantly, we also directly compare the performance for upright faces in the sham vs anodal tDCS group. This comparison is also motivated by previous work conducted in our lab (Civile et al., 2016, and more recent pilot work by Civile, Obhi, McLaren, 2018) where anodal tDCS delivered over the Fp3 was found to reduce performance for upright familiar checkerboards or faces, compared to the same condition in the sham group. This comparison is particularly appropriate because the same stimulus sets are rotated across participants in a counterbalanced manner; so that each upright face seen in the anodal group for a given participant will equally often serve as an upright face for the participants in the sham group.

Experiment 1 (Fp3). A 2 x 2 ANOVA produced a significant interaction, $F(1, 46) = 5.93, p = .019, \eta^2_p = .11$, indicating that the inversion effect differed between the two groups. A significant inversion effect was found in the sham group, $t(23) = 6.12, p < .001, \eta^2_p = .62$, and, critically, a *reduced* (but still significant) *inversion effect* was found in the tDCS anodal group, $t(23) = 3.50, p = .002, \eta^2_p = .34$ (see Figure 2). An additional analysis showed that recognition for upright face stimuli in the anodal group was significantly reduced compared to that in the sham group, $t(46) = 2.62, p = .015, \eta^2_p = .24$. No difference was found for inverted faces in the two groups, $t(46) = .50, p = .619, \eta^2_p = .01$.

Experiment 2 (Replication). As in Experiment 1 we found a significant interaction, $F(1, 46) = 5.93, p = .006, \eta^2_p = .15$. Thus, a significant inversion effect was found in the sham

group, $t(23) = 7.66, p < .001, \eta^2_p = .71$, and once again we found a significantly reduced inversion effect in the tDCS anodal group, $t(23) = 2.91, p = .008, \eta^2_p = .27$ (see Figure 2). In this experiment, as in Experiment 1, recognition for upright face stimuli in the anodal group was significantly reduced compared to that in the sham group, $t(46) = 3.18, p = .004, \eta^2_p = .26$. No difference was found for inverted faces in the two groups, $t(46) = .05, p = .957, \eta^2_p = 0$.

Experiment 3 (Active Control). A 2 x 2 ANOVA produced no reliable interaction, $F(1, 46) = 0.65, p = .42, \eta^2_p = .01$. A significant inversion effect was found in the sham group, $t(23) = 4.04, p < .001, \eta^2_p = .41$, and in the tDCS anodal group, $t(23) = 5.51, p < .001, \eta^2_p = .57$, i.e. this montage did not result in the reduced inversion effect found in Experiments 1 and 2 (see Figure 2).

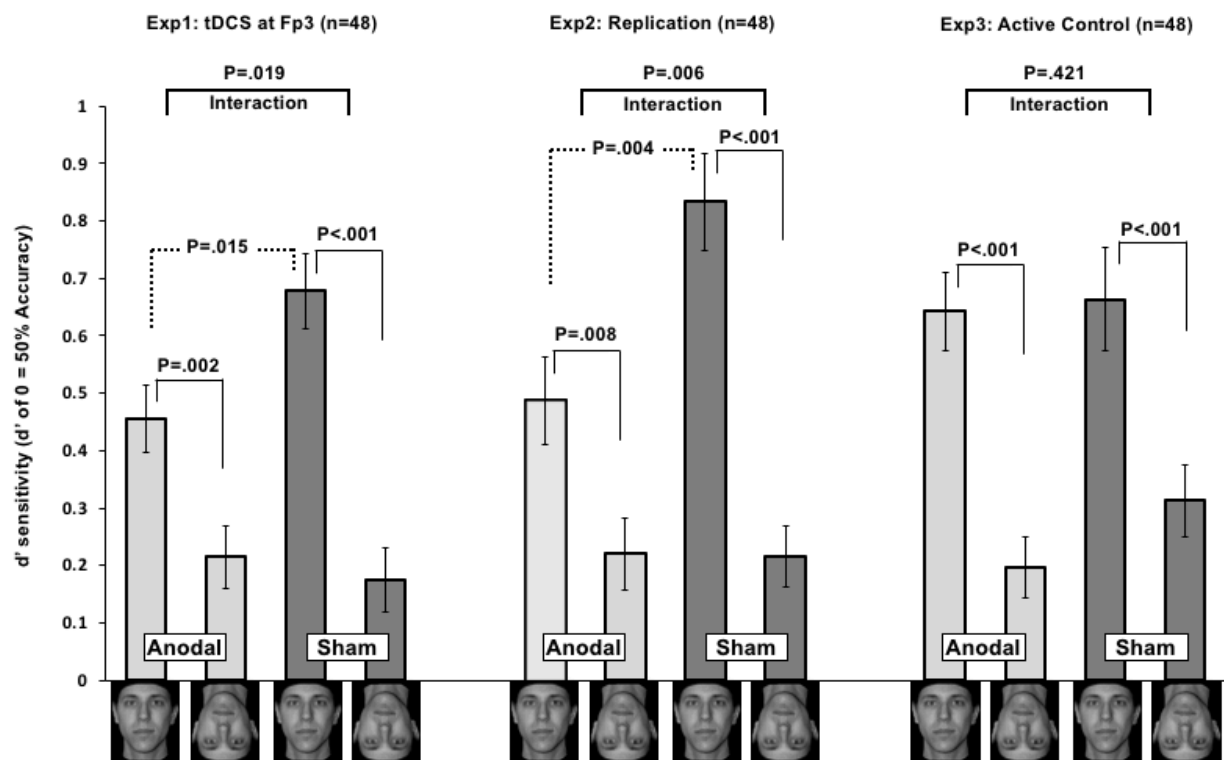


Figure 2 reports the results from the three experiments. The x-axis shows the stimulus conditions. The y-axis shows sensitivity d' measure. Error bars represent s.e.m.

Bayes Factor Analysis

Using the procedure outlined by Dienes (2011), we first conducted a Bayes analysis on the Face Orientation by Stimulation interaction found in both Experiment 1 and 2. Thus, we used the interaction found in Experiment 1 as the prior, setting the standard deviation of p (population value | theory) to the mean for the difference between the inversion effect in sham group minus that in the anodal group (0.26). We used the standard error (0.10) and mean difference (0.35) between the inversion effect in the sham group minus that in the anodal group in Experiment 2 in our calculation. We assumed a one-tailed distribution for our theory and a mean of 0. This gave a Bayes factor of 148, which is very strong evidence indeed for the theory (because greater than 3, for the conventional cut-offs see Jeffrey et al., 1961), in this case that the interaction will be positive and non-zero. Because in both Experiment 1 & 2 performance for the upright faces was significantly better in the sham group compared to that in the anodal group, we calculated the Bayes factor for this effect using as prior Experiment 1, setting the standard deviation of p as the mean difference between sham upright faces minus anodal upright faces (0.22). We then used the standard error (0.11) and mean difference (0.34) between sham upright faces minus anodal upright faces in Experiment 2 in the calculation. Once again, we assumed a one-tailed distribution for our theory and mean of 0. This gave a Bayes factor of 40, which is again very strong evidence that the performance on the sham upright faces will be higher than that on the upright faces experienced under anodal tDCS. Finally, we calculated the Bayes factor for the Face Orientation by Stimulation interaction in Experiment 3 using as the prior (standard deviation of p) the interaction averaged over Experiment 1 and 2 (0.30). We used the standard error (0.13) and mean difference (-0.09) for the interaction in Experiment 3, and again assumed a one-tailed distribution for our theory and a mean of 0. This gave a Bayes factor of 0.26, which is good evidence for the null (because close to 0).

Discussion

Building on previous work with checkerboards (Civile et al., 2016) and pilot work with faces (Civile, Obhi and McLaren, 2018), the evidence we present here establishes that anodal tDCS to Fp3 using our electrode montage (based on that used by Ambrus et al, 2011) affects the mechanisms supporting perceptual learning in humans; and this impairs their ability to discriminate between faces, mainly by reducing the advantage enjoyed by upright relative to inverted faces in a standard study/test recognition experiment. But we will argue that it is not just a matter of "switching off" perceptual learning as might be thought by taking Civile et al (2016)'s title at face value. Rather, it can be described as a reconfiguration of the processing that produces representations of stimuli, such that instead of pre-exposure to a prototype-defined category enhancing the discriminability of (even novel) exemplars taken from that category (McLaren, Leivers and Mackintosh, 1994; McLaren, 1997), it now enhances generalization between them, and makes features common to those exemplars more prominent rather than exaggerating their differences. As a consequence, the stimulation can be said to change how the mind works, rather than just disabling a component of it. This is a bold claim, and we present our arguments in support of it below.

In Experiment 1, we show that anodal tDCS delivered (10 mins at 1.5mA) over the left DLPFC at Fp3 significantly reduces the face inversion effect relative to a sham control condition under double blind conditions. This is the first demonstration of this kind. Experiment 2 confirmed this effect by replicating the pattern of results found in Experiment 1 to establish its reliability. These results are supported by the Bayesian analysis that we conducted on the critical *Face Orientation by Stimulation* interaction using priors from Experiment 1, which gives a Bayes factor of 148, providing conclusive evidence in favour of the hypothesis that the inversion effect is greater in sham than in anodal groups using this procedure. In both Experiment 1 and 2 performances for upright faces in the anodal group

was impaired compared to sham. Finally, in Experiment 3, we show that stimulation applied to a different targeted brain area i.e. rIFG (our active control) does not have the same effect, providing the first evidence that the reduction in the inversion effect is not obtained by stimulating any brain area. Taking all three studies together, we suggest that the reduction in the inversion effect, in those conditions in which it occurs, is due mainly to a reduction in performance on the upright faces consequent on Fp3 stimulation, with the inverted faces relatively unaffected. This last point establishes that the effect is not merely due to a general decrement in learning, making the case for an effect based on perceptual learning *per se*.

To see this, consider that the main finding emerging from these experiments is that the robust inversion effect that would usually be obtained for face stimuli is reduced when relatively brief tDCS stimulation is delivered over the Fp3 brain area. We are arguing that this cannot be due to tDCS affecting recognition performance in general, because if it was, then the anodal groups in Experiments 1 and 2 would have shown a reduction in recognition performance (compared to sham) for inverted faces as well as for upright faces. However, this was not the case, and inverted faces were recognised significantly above chance and, most importantly, showed no differences across groups.

To accommodate this pattern of results we propose that tDCS has a quite specific effect on perceptual learning, in that it disrupts the modulation of feature salience based on prediction error that normally produces perceptual learning in these circumstances (McLaren, et al, 2016; Civile et al, 2016). It is this change in perceptual learning that causes the reduction in the face inversion effect, because it reduces people's ability to discriminate between different upright faces, which is normally enhanced by their expertise for face processing acquired via experience and manifesting as perceptual learning. Under normal conditions, the theory predicts that the salience of units representing features of the stimulus that have strong associations with other units that are also active is reduced by this process

(see any of McLaren, Kaye and Mackintosh, 1989; McLaren and Mackintosh, 2000; McLaren, Forrest and McLaren, 2012 for a discussion). This has the effect of making the prototypical features of an exemplar (which it will tend to share with other exemplars drawn from that category) less salient and enhances the relative salience of features unique to that exemplar, thus improving its discriminability from other exemplars. Disrupting this process via tDCS means that the same associations that led to a low error and hence low salience for some features now make these units more (rather than less) active and quite possibly the most salient. This will now enhance *generalization* between exemplars as a consequence of familiarity with that category, rather than producing the enhanced discriminability that is the hallmark of perceptual learning (McLaren, et al, 2016). The result is the elimination of the inversion effect seen in Civile et al (2016)'s study with checkerboards, and the reduction of the inversion effect for faces reported here.

One reason why we believe this is the correct explanation for the effects reported in this paper is that, in contrast to other possible theories of the basic effect on perceptual learning, this account has the advantage of explaining why tDCS can affect the results of perceptual learning acquired over a lifetime of experience (e.g. with faces), rather than just the perceptual learning process itself. To see this, recall that our demonstration with checkerboards used tDCS during what was, effectively, the stimulus exposure phase. The elimination of any perceptual learning effect in those circumstances could have been due to tDCS simply preventing perceptual learning from occurring. But now we have a result where an effect, at least partially attributable to perceptual learning (the inversion effect for faces) but dependent on perceptual learning that has taken place over a long period before stimulation was applied, still produces a similar result. This quite clearly indicates that it is not the process by which exposure leads to perceptual learning that must necessarily be impacted to produce our results. It is possible to transform processing of the stimuli in a way

that is based on learning that has taken place much earlier to arrive at our effect. Changing the way that salience is modulated based on prediction error (i.e. past learning) fits this specification. Other theories that rely on the allocation of attention or some process of comparison (e.g. Mundy, Dwyer and Honey, 2006; Mundy, Honey and Dwyer 2007, Wang, Lavis, Hall and Mitchell, 2012; Jones and Dwyer, 2013) do not so obviously do so.

One may notice that, while anodal tDCS over the Fp3 significantly reduced the face inversion effect compared to sham, it did not eliminate it entirely. This brings us to one of the main debates about face recognition: *Are faces special or not?* Clearly, our results must offer some support to the expertise account of face recognition, but if the face inversion effect was entirely due to perceptual learning, then, by analogy with the results obtained with the checkerboard stimuli in Civile et al (2016)'s work, we would have expected to eliminate the inversion effect for our face stimuli. However, this was not the case. One potential explanation is that there may be two separate components contributing to the face inversion effect. One is that due to perceptual learning, a component that we have shown we are able to influence by means of tDCS stimulation. But, this leaves the possibility that the other component may well be specific to face stimuli. The current data do not directly address this issue, but future studies should. There are, of course, other potential explanations for the residual inversion effect observed with faces even under stimulation using our procedures, and we note them here as they may provide a challenge to the MKM-based approach that we have adopted. It may be that the stimulation itself was not strong enough to completely eliminate the effect, but recall that it was for the checkerboards. We can try to explain this difference by appealing to the fact that people have so much experience of faces, and this could make the perceptual learning based on that experience particularly resistant to modification. But there is a potential difficulty with this argument. To succeed, it requires some reduction in the benefits of perceptual learning that is not complete. How can this be achieved? If we take the MKM

approach as an example, we can imagine reducing the error-based modulation of salience just enough to get the desired effect. But this reduction would be in no way dependent on the amount of experience with the stimuli, it is an effect on the mechanism for perceptual learning itself, and so should also have manifested in the case of the checkerboards. Clearly, then, our MKM-based approach cannot easily explain the results with faces and those with checkerboards without appealing to some other process at work in the case of faces. But the problem here, if we dispense with the MKM account, is to find a theory of perceptual learning that would naturally integrate amount of experience with a stimulus class with the effects of tDCS, and at present, there are no obvious candidates. This is obviously an area for further research.

To sum up, in this paper we have shown how a brief anodal tDCS stimulation at Fp3 on the scalp is able to reduce the inversion effect for faces, an effect analogous to previous effects found for checkerboard stimuli drawn from familiar categories. Our current results contribute to the perceptual learning account of face recognition, and provide evidence for a causal relationship between tDCS delivered in a manner that affects perceptual learning and the reduction of the inversion effect that would otherwise be obtained for familiar checkerboards and for face stimuli.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 743702 awarded to Ciro Civile. This project has also received funding from the Economic and Social Research Council (ESRC) New Investigator Grant (Ref. ES/R005532) awarded to Ciro Civile (PI) and I.P.L. McLaren (Co-I).

References

- Ambrus G. G., Zimmer M., Kincses Z. T., Harza I., Kovacs G., Paulus W., et al. (2011). The enhancement of cortical excitability over the DLPFC before and during training impairs categorization in the prototype distortion task. *Neuropsychologia* 49, 1974–1980.
- Carmel, D., and Bentin, S. (2002). Domain specificity versus expertise: factors influencing distinct processing of faces. *Cognition*, 83, 1-29.
- Civile, C., Zhao, D., Ku, Y., Elchlepp, H., Lavric, A., and McLaren, I.P.L. (2014). Perceptual learning and inversion effects: Recognition of prototype-defined familiar checkerboards. *Journal of Experimental Psychology: Animal Behavior Processes*, 40, 144-61.
- Civile, C., Verbruggen, F., McLaren, R., Zhao, D., Ku, Y., and McLaren, I.P.L. (2016). Switching off perceptual learning: Anodal transcranial direct current stimulation (tDCS) at Fp3 eliminates perceptual learning in humans. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 290-296.
- Civile, C., McLaren, R., and McLaren, I.P.L. (2016). The face inversion effect: Roles of first and second-order relational information. *The American Journal of Psychology*, 129, 23-35.
- Civile, C., McLaren, R., and McLaren, I. P. L. (2014). The face inversion effect: Parts and wholes. *The Quarterly Journal of Experimental Psychology*, 67, 728–746.
- Civile, C., Elchlepp, H., McLaren, R., Galang, C.M., Lavric, A., and McLaren, I.P.L. (2018). The effect of scrambling upright and inverted faces on the N170. *Quarterly Journal of Experimental Psychology*, doi: 1747021817744455.
- Civile, C., Obhi, S.S., and McLaren, I.P.L. (2018). The Role of Experience-Based Perceptual Learning in the Face Inversion Effect. *Vision Research*, doi.org/10.1016/j.visres.2018.02.010
- Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). HillSEale, NJ: Lawrence Earlbaum Associates.

Cunillera, T., Brignani, D., Cucurell, D., Fuentemilla, L., & Miniussi, C. (2016). The right inferior frontal cortex in response inhibition: A tDCS-ERP co-registration study.

NeuroImage, 140, 66–75.

Cunillera, T., Fuentemilla, L., Brignani, D., Cucurell, D., & Miniussi, C. (2014). A simultaneous modulation of reactive and proactive inhibition processes by anodal tDCS on the right inferior frontal cortex. *PLoS ONE*, 9, e113537.

Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107-117.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6, 274–290.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6, 274–290.

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, 11, 2319-2324.

Gauthier, I., and Tarr, M. (1997). Becoming a “Greeble” expert: exploring mechanisms for face recognition. *Vision Research*, 37, 1673-1682.

Gold, J., & Watanabe, T.(2010). Perceptual learning. *Current Biology*, 20, R46–48.

Jacobson, L., Javitt, D. C., & Lavidor, M. (2011). Activation of inhibition: Diminishing impulsive behavior by direct current stimulation over the inferior frontal gyrus. *Journal of Cognitive Neuroscience*, 23, 3380–3387.

Jeffreys H.(1939/1961). *The Theory of Probability* 1st/3rd Edn. Oxford, England: Oxford University Press

Jones, S. P., & Dwyer, D. M. (2013). Perceptual learning with complex visual stimuli is based on location, rather than content, of discriminating features. *Journal of Experimental Psychology: Animal Behavior Processes*, 39, 152–165.

McLaren, I.P.L. (1997). Categorization and perceptual learning: An analogue of the face inversion effect. *The Quarterly Journal of Experimental Psychology* 50A, 257-273.

McLaren, I.P.L., Carpenter, K., Civile, C., McLaren, R., Zhao, D., Ku, Y., Milton, F., and Verbruggen, F. (2016). Categorisation and Perceptual Learning: Why tDCS to Left DLPC enhances generalisation. In *Associative Learning and Cognition*. Homage to Prof. N.J.

Mackintosh. Trobalon, J.B., and Chamizo, V.D. (Eds.), University of Barcelona.

McLaren, I. P. L., Leavers, H. L., & Mackintosh, N. J. (1994). Recognition, categorisation and perceptual learning. In C. Umiltà & M. Moscovitch (Eds.), *Attention & performance XV* (pp. 889–909). Cambridge, MA: MIT Press.

McLaren, I.P.L., Kaye, H. & Mackintosh, N.J. (1989). An associative theory of the representation of stimuli: Applications to perceptual learning and latent inhibition. In R.G.M. Morris (Ed.) *Parallel Distributed Processing - Implications for Psychology and Neurobiology*. Oxford, Oxford University Press.

McLaren, I.P.L. and Mackintosh, N.J. (2000). An elemental model of associative learning: Latent inhibition and perceptual learning. *Animal Learning and Behavior*, 38, 211-246.

McLaren, I.P.L., Forrest, C.L., McLaren, R.P. (2012). Elemental representation and configural mappings: combining elemental and configural theories of associative learning. *Learning and Behavior*, 40, 320-333.

McLaren, I. P. L., Wills, A. J., & Graham, S. (2010). Attention and perceptual learning. In C. Mitchell & M. Le Pelley (Eds.), *Attention and associative learning: From brain to behaviour* (pp. 131–158). Oxford, England: Oxford University Press.

Mundy, M. E., Dwyer, D. M., & Honey, R. C. (2006). Inhibitory associations contribute to perceptual learning in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, 32, 178–184.

- Mundy, M. E., Honey, R. C., & Dwyer, D. M. (2007). Simultaneous presentation of similar stimuli produces perceptual learning in human picture processing. *Journal of Experimental Psychology: Animal Behavior Processes*, 33, 124–138.
- Stramaccia, D. F., Penolazzi, B., Sartori, G., Braga, M., Mondini, S., & Galfano, G. (2015). Assessing the effects of tDCS over a delayed response inhibition task by targeting the right Inferior Frontal Gyrus and right Dorsolateral Prefrontal Cortex. *Experimental Brain Research*, 233, 2283–2290.
- Valentine, T., & Bruce, V. (1988). Mental rotation of faces. *Memory & Cognition*, 16, 556–566.
- Wang, T., Lavis, Y., Hall, G., & Mitchell, C. J. (2012). Location and salience of unique features in human perceptual learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 38, 407–418.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141–145.
- Yovel G., Kanwisher N. (2005) The neural basis of the behavioral face-inversion effect *Current Biology*, 15(24) 2256- 62.
- Zion-Golumbic, E., & Bentin, S. (2007). Dissociated neural mechanisms for face detection and configural encoding: Evidence from N170 and Gamma-band oscillation effects. *Cerebral Cortex*, 17, 1741–1749.